doi. 10. 3724/1005-0566. 20251003

人工智能基础设施的关键要素与建设路径研究

隆云滔[1,2],黄云游[3],许哲平[4,5],王敬敬[6,7],王　哲[8],任旭东[9]

(1. 中国科学院科技战略咨询研究院,北京　100190; 2. 中国科学院大学公共政策与管理学院,北京　100049;
3. 中国社会科学院大学国际政治经济学院,北京　102488;4. 中国科学院文献情报中心,北京　100190;
5. 中国科学院大学经济与管理学院信息资源管理系,北京　100190;6. 北京信息科技大学商学院,北京　102206;
7. 北京市知识管理研究基地,北京　102206;8. 北京航空航天大学公共管理学院,北京　100191;
9. 复杂关键软件环境全国重点实验室,北京　100191)

摘　要:在数字经济时代,人工智能基础设施已经成为人工智能发展的核心依托,正在推动全球人工智能产业生态重塑、大国竞争场域拓展。本文着眼于人工智能基础设施建设的关键要素研究,从基础层、模型层、应用层、政策层系统梳理全球发达国家部署、运用人工智能基础设施的实践经验与政策举措,分析我国在人工智能基础设施建设方面的挑战,为我国在人工智能领域提升创新效能、明确能力边界、规制技术风险提供理论和现实依据。基于此,围绕基础技术、模型架构、应用融合、政策保障提出政策建议。

关键词:人工智能;数字基础设施;开放科学;数字化;国际经验

中图分类号:TP18　　**文献标识码:**A　　**文章编号:**1005 - 0566(2025)10 - 0027 - 12

Research on the key elements and development paths of AI infrastructure

LONG Yuntao[1,2], HUANG Yunyou[3], XU Zheping[4,5], WANG Jingjing[6,7], WANG Zhe[8], REN Xudong[9]

(1. *Institutes of Science and Development*, *Chinese Academy of Sciences*, *Beijing* 100190, *China*;
2. *School of Public Policy and Management*, *University of Chinese Academy of Sciences*, *Beijing* 100049, *China*;
3. *School of International Politics and Economics*, *University of Chinese Academy of Social Sciences*, *Beijing* 102488, *China*;
4. *National Science Library*, *Chinese Academy of Sciences*, *Beijing* 100190, *China*;
5. *Department of Information Resources Management*, *School of Economics and Management*, *University of Chinese Academy of Sciences*, *Beijing* 100190, *China*;
6. *Business School*, *Beijing Information Science and Technology University*, *Beijing* 102206, *China*;
7. *Beijing Knowledge Management Research Center*, *Beijing* 102206, *China*;
8. *School of Public Administration*, *Beihang University*, *Beijing* 100191, *China*;
9. *State Key Laboratory of Complex and Critical Software Environment*, *Beijing* 100191, *China*)

Abstract: In the era of digital economy, artificial intelligence infrastructure has become the core support for the development of artificial intelligence, driving the reshaping of the global AI industrial ecosystem and the expansion of the competition arena among major countries. This article focuses on the study of key elements of AI infrastructure construction, systematically reviews the practical experiences and policy measures of developed countries in deploying

收稿日期:2025-06-29　　**修回日期:**2025-09-22

基金项目:国家社会科学青年基金“大模型开源开放创新生态发展的治理机制研究”(24CJY018)。

作者简介:隆云滔(1984—),女,湖南邵阳人,中国科学院科技战略咨询研究院副研究员,经济学博士,研究方向为开源创新发展、人工智能开源与治理、复杂经济学。通信作者:任旭东。

and applying AI infrastructure from the perspective of the foundation layer, model layer, application layer, and policy layer, and analyzes the challenges China faces in AI infrastructure construction. It provides theoretical and practical basis for China to enhance innovation efficiency, clarify capability boundaries, and regulate technological risks in the field of artificial intelligence. Finally, the article puts forward policy suggestions around basic technologies, model architectures, application integration, and policy guarantees.

Key words: artificial intelligence; digital infrastructure; open science; digitalization; international experience

基础设施作为推动人工智能(artificial intelligence, AI)发展的关键支撑,其战略地位日益突显。我国高度重视人工智能基础设施的建设。2017 年 7 月,国务院发布《新一代人工智能发展规划》,开启了我国人工智能发展的长远布局。2024 年 9 月,工业和信息化部等 11 个部门共同发布《关于推动新型信息基础设施协调发展有关事项的通知》,对新型基础设施的跨领域、跨行业、跨网络建设进行全方位战略部署。2025 年 5 月,国家数据局发布《数字中国建设 2025 年行动方案》,将“人工智能+”、数字基础设施提质扩容作为未来行动的重点。2025 年 7 月,在 2025 世界人工智能大会暨人工智能全球治理高级别会议上发布《人工智能全球治理行动计划》,提出让人工智能成为造福人类的国际公共产品。2025 年 8 月,《国务院关于深入实施“人工智能+”行动的意见》提出“推动人工智能普惠共享,把人工智能作为造福人类的国际公共产品”。尽管如此,我国要精准把握数字时代脉搏、成为全球人工智能产业生态中的关键角色,仍然需要进一步强化人工智能基础设施部署、加快人工智能基础设施技术攻关、完善人工智能基础设施资源统筹。以美国为代表的发达国家在人工智能基础设施建设中已积累丰富的实践经验。本文将系统梳理发达国家在人工智能基础设施建设中的先进做法,剖析我国人工智能基础设施建设面临的主要问题,为我国打造具备国际领先水平的人工智能基础设施提供参考借鉴。

一、人工智能基础设施的发展现状

(一)人工智能基础设施:概念定义与历史演进

当前,政府和产业界均对人工智能基础设施进行了界定。工业和信息化部等部门将人工智能基础设施定义为人工智能技术推广普及过程中形成的一类新型基础设施,包括支持开发的人工智能算法框架、面向应用的人工智能算法平台和公共服务平台、用于算法模型训练的公共数据集等形态[1]。开源软件公司红帽(Red Hat)将人工智能基础设施定义为:结合人工智能和机器学习(AI/ML)技术,用于开发和部署可靠、可扩展的数据解决方案。它的技术栈包括 3 层:基础设施层(构建和训练模型所需的硬件和软件)、应用层和模型层(通用人工智能、特定人工智能和超本地人工智能)[2]。高通和戴尔集团联合发布的报告称,人工智能基础设施是一个相互依存的生态系统,包括数据平台、人工智能模型、数据中心硬件、网络、半导体和存储[3]。综合而言,人工智能基础设施是一个集成硬件、软件、数据与算法的复杂系统,其根本目标在于支撑 AI 研发与应用的全生命周期。其中,政府的定义更强调人工智能基础设施公共产品的属性,产业界的定义则侧重于人工智能基础设施的技术属性和商业应用。鉴于对人工智能基础设施的全面考察需要同时兼顾技术内涵与政策导向,本文从基础层、模型层、应用层与政策层 4 个维度理解人工智能基础设施。其中,基础层是指算力、算法、数据和电力能源等底层支撑要素,模型层包括作为通用能力底座的人工智能大模型及其研发生态,应用层涵盖人工智能技术与不同行业领域的融合与应用场景,政策层是指人工智能基础设施规划、治理和监管的顶层设计。

人工智能基础设施的历史演进与技术发展密不可分。20 世纪 50 年代,达特茅斯学院(Dartmouth College)举办的科学计算研讨会首次提出“机器模拟人类智能”这一设想,认为人工智能是可实现的概念。这一时期,人工智能基础设施主要作为概念验证的实验平台存在,大型计算机构成了人工智能研究的主要依托,但其在计算能力、储存容

量、算法上都受到极大的限制。匮乏的技术基础使得人工智能研究在此后几十年陷入停滞状态。20 世纪 80 年代,机器学习的兴起标志着人工智能基础设施的重大转向,对于计算能力需求的指数级增长催生了硬件和软件的进步,奠定了人工智能研究的技术基础。英伟达(NVIDIA)于 1999 年推出首款真正意义上的图形处理器(GPU),人工智能计算硬件逐渐从通用性向专用性演化。21 世纪初,云计算与大数据的兴起进一步推动了人工智能基础设施的发展。亚马逊网络服务(AWS)、谷歌云和微软 Azure 等云平台提供了可扩展的计算资源,是云人工智能基础设施的有力支撑。2016 年,谷歌推出了专为人工智能训练设计的张量处理单元(TPU),其强大的并行计算能力为训练更复杂的神经网络和处理海量数据提供了关键的硬件支持。近年来,以英伟达和华为等为代表的企业,通过发展大规模 GPU 集群与高速互联技术,将算力规模推向新的高度。当前,以深度学习为核心的人工智能正处在快速发展阶段,算力基础设施已成为其进步的核心驱动力。在此背景下,边缘计算与人工智能的深度融合,正成为塑造未来产业格局的关键方向[4]。

(二)人工智能基础设施建设的国家间差异

鉴于人工智能基础设施与人工智能发展之间的强关联性,全球相关机构开展的人工智能国家排名工作在一定程度上揭示了不同国家人工智能基础设施的发展水平。斯坦福大学 2024 年 11 月发布的《全球人工智能活力工具》主要探讨全球人工智能活力工具(GVT)的设计和实施,从研发、负责任 AI、经济、教育、多样性、政策治理、公众意见和基础设施等 8 个维度[5],通过 42 项指标对 36 个国家的人工智能活跃度进行综合评估和排名(见表 1)。在全球人工智能(AI)发展的综合排名中,国家竞争格局可分为 3 个梯队。第一梯队(主导层):由美国和中国领跑。第二梯队(稳定层):英国和印度。第三梯队(波动层):涵盖法国、德国、日本、新加坡和韩国等国家。

表 1　《全球人工智能实力排行榜》的 TOP 4 国家详细得分的对比表

国家	排名	总得分	研发	负责任 AI	经济	教育	多样性	政策治理	公众意见	基础设施
美国	1	70.06	19.19	5.71	13.55	2.75	1.01	6.94	3.99	16.91
中国	2	40.17	14.78	1.96	6.19	0.94	1.08	4.40	1.33	9.49
英国	3	27.21	2.58	1.11	5.92	3.69	1.25	9.67	0.65	2.35
印度	4	25.54	9.37	0.26	4.30	2.02	2.86	4.55	1.58	0.60

《2025 年敏捷指数报告》对全球人工基础设施的建设情况进行详细对比[6]。报告从数据中心数量、云服务支出、宽带覆盖率与计算能力 4 个方面衡量人工智能基础设施的实力。其中,美国在基础设施维度上得分最高(16.91 分),中国位居第二(9.49 分),显示了强大的增长实力,但与美国仍存在较大差距。其他国家如英国(2.35 分)和印度(0.60 分)在这一维度上的得分则远低于美中两国。就美中对比而言,美国的主要优势在于数据中心和云服务方面,美国拥有 2 307 个数据中心,在数据中心数量上占据绝对优势。此外,美国在云基础设施服务支出方面同样处于主导地位,其支出金额是中国的 4 倍之多。

作为人工智能基础设施的重要组成部分,大语言模型的优势和排名在一定程度上也反映了各国在人工智能基础设施上的综合实力。国际权威评测平台 Artificial Analysis 利用智能指数对全球 2025 年第一季度的主要大语言模型进行评测和排名[7]。在得分 55 分以上的 21 个可评测大模型中,美国的大模型数量为 11 个,占了一半,包括 OpenAI o3、OpenAI o1、OpenAI o3-mini、OpenAI o1-mini、Gemini 2.0 Flash、Gemini 1.5 Pro、Claude 3.5、GPT-4o、Nove Pro、Llama 3.3、Grok 3.0 Beta。中国有 7 个,法国、加拿大和以色列各有 1 个。

(三)人工智能基础设施的研究网络文献分析

在 Web of Science(WoS)的 Core Collection(核心合集)数据库上,限定年份为近 5 年(2020—2024 年),采用主题词“AI infrastructure”进行检索,共检索到 3 343 条结果,检索日期为 2025 年 3 月 31 日。对检索的论文进行研究主题分析,以 Tab Delimited File 格式下载 Full Record(完整记录)内容,利用 VOS Viewer 文献计量分析工具对论文关键词进行可视化图谱统计分析。由此可以发现,人工智能基础设施相关的研究主题主要包括以下 6 个方面:一是 AI 基础设施的技术架构与平台,包括云计算、边缘计算、分布式计算等;二是 AI

算法与模型创新,专注于提高 AI 系统的性能、效率和可解释性;三是 AI 与物联网(IoT)的融合,如智能家居、智能城市、工业物联网、数字孪生和区块链等;四是 AI 在基础设施中的应用与优化,如智能交通系统、智能电网、智能建筑等;五是 AI 在不同行业领域的应用,包括医疗健康、教育学习和农业与自然资源管理、可持续发展与绿色技术领域等;六是 AI 安全、伦理与治理,包括数据隐私、算法偏见、对抗性攻击、算法透明度等。

二、人工智能基础设施建设的国际经验

人工智能基础设施极具战略性,欧美发达国家及相关政府和组织通过系统化的政策措施,围绕基础技术、模型架构、应用融合、政策保障 4 个维度进行全方位布局,打造有效的人工智能基础设施建设生态,推动人工智能基础设施实现稳步发展与有序部署。

(一)基础层:夯实人工智能基础设施的内、外部核心支撑

人工智能基础设施基础层主要涉及内部的算力、算法、数据,以及外部的电力能源,基础层的建设构成了人工智能基础设施的核心支撑。其主要包括以下 4 个方面。

(1)整合算力、算法与数据资源,打造国家级基础设施平台。美国、欧盟等基于国家力量整合算力、算法与数据,以构建高效的计算与人工智能资源支持生态系统。美国的国家人工智能研究资源(NAIRR)旨在通过对研究人员开放高性能计算、数据集、软件、模型等资源,加速人工智能领域的创新发展和技术突破。其包括 4 个重点领域,分别是开放人工智能资源的 NAIRR Open、专注隐私数据安全的 NAIRR Secure、促进软件平台互联互通的 NAIRR Software 开展人工智能培训的 NAIRR Classroom。预计 NAIRR 建成后的算力资源能够支持至少 1 万亿参数规模机器学习模型训练的超级计算机,以及云计算资源、CPU、GPU 和高速网络。欧盟下一代云基础设施和服务(IPCEI CIS)项目由 19 个高度创新的子项目组成[8],意在建立一个开放且可互操作的欧洲数据处理生态系统,涉及基础软件、硬件乃至特定行业应用的完整技术体系。

(2)布局核心算力、算法设施。算力、算法设施技术要求高、资金投入多、建设周期长,美国对此采取多部门协同推进的战略。美国对超级计算机的系统性支持始于 20 世纪 80 年代,美国国家科学基金会(NSF)、能源部(DOE)、国家航空航天局(NASA),以及国防高级研究计划局(DARPA)当时共同启动了大规模的计算计划。其中,NSF 主要负责资助建立全国性超算中心网络及相应的数据网络,DOE 聚焦于超级计算机的使用和研发,NASA 致力于处理超级计算机与航空航天任务的融合运用,DARPA 则为前沿和高风险的计算机体系结构研究提供支持[9]。这一支持体系构成了美国超级计算机发展的基础,推动美国超级计算机在数量与质量上迅速发展。截至 2024 年,美国已拥有 173 台全球前 500 的超级计算机,聚合算力世界排名第一[10]。到 2025 年,全球 AI 基础设施支出预计将超过 2 000 亿美元,云提供商和硬件公司将投资超过 3 200 亿美元用于数据中心和 AI 专用硬件。为应对美国数据中心产业的迅猛扩张,并引导其迈向更具韧性、安全与可持续的发展路径,美国国家电信和信息管理局(NTIA)于 2024 年 5 月正式发布《关于加强数据中心发展、韧性与安全》(Request For Comments on Bolstering Data Center Growth, Resilience, and Security)的征求意见稿。该倡议旨在通过广泛的公众咨询,系统性地应对数据中心发展面临的多维挑战。其核心议题主要包括 6 个维度:①在市场与规制层面,识别并消除制约数据中心投资与建设的市场壁垒与监管障碍;②在能源与基础设施层面,保障稳定、充裕的电力供应,并推动能源结构的优化;③在安全与风险管理层面,审查当前的数据中心物理与网络安全实践及标准;④在人力资源与劳动力层面,应对数据中心设计、建设与运维领域的专业人才短缺问题;⑤在供应链可靠性层面,评估关键信息技术与运营技术设备所面临的供应链风险;⑥在环境与社会影响层面,分析数据中心扩张对当地社区、资源消耗及环境的综合影响。

(3)发展绿色人工智能基础设施。人工智能基础设施建设必然带来相应的碳成本[11]。美国的

大型科技企业致力于通过动态能源管理和能效技术优化其数据中心的能源使用。此外,清洁能源采购和碳抵消也成为降低碳足迹的重要方式。一方面,在投资太阳能与风能技术的同时,微软正探索小型模块化反应堆(SMR)与传统核电设施的混合部署模式,以满足数据中心和人工智能日益增长的需求。另一方面,微软聚焦于利用可持续材料在建筑项目中打造低碳云基础设施,并在具体运营中利用内部碳费机制来确保运营过程实现碳中和。

(4)建立有效的碳排放报告机制。例如,欧盟委员会授权条例(EU)2024/1364 引入新的数据中心可持续性报告义务,要求 500 kW 以上的数据中心运营商必须报告建筑面积、安装功率、数据量、能源消耗、PUE、温度设定点、废热利用,以及水和可再生能源使用情况[12]。美国《2024 年人工智能环境影响法案》(Artificial Intelligence Environmental Impacts Act of 2024)要求环境保护署署长对人工智能的环境影响进行全面研究,涵盖人工智能模型和硬件的整个生命周期,同时强调要建立人工智能环境影响实体自愿报告制度、建立专门的人工智能环境影响联盟。

(二)模型层:构建完善的大模型研发生态

美国是大模型研发方面的领导者。2024 年,美国机构共发布 40 个知名的人工智能大模型,其中谷歌、OpenAI、Meta 和微软是模型研发的核心主体[13]。这一成就有赖于美国国内完善的大模型研发生态。

(1)成熟的开源平台。开源平台正在显著提升大模型领域的研发效率并驱动创新。其中,美国的 Hugging Face 已发展为全球最具影响力的机器学习社区与协作平台。该平台的核心架构由四大关键要素构成:模型(Models)、数据集(Datasets)、代码仓库(Repositories)与应用空间(Spaces)。其核心贡献在于,通过为模型与数据集提供基于 Git 的版本控制和协作机制,构建了一套标准化的机器学习资产管理工作流。这使得开发者能够像管理代码一样,便捷地托管、版本化、共享与迭代机器学习模型与数据集,极大地促进了知识的沉淀与项目的复现。模型是 Hugging Face 的核心部分,平台已托管超过 38 万预训练模型,涵盖自然语言处理、计算机视觉、语音识别和多模态等多个领域。同时,平台已汇聚超过 40 万个数据集,为模型研发提供关键支撑。应用空间则允许开发者将模型快速部署为可交互的演示界面,便于研究成果在社区中传播与验证。

(2)打造全栈技术生态链。美国构建了从底层硬件到高层应用的全栈技术生态链,通过贯通算力芯片、操作系统、算法框架、开发工具到行业应用的全链条布局,形成强大的技术链协同效应。其中具有代表性的是 NVIDIA 的 CUDA 生态。CUDA 生态一共包含 6 层技术架构,分别是硬件层(NVIDIA GPU 与系统)、核心软件层(CUDA 平台与 Magnum IO)、开发层(编程方法与工具)、软件库层(CUDA-X SDKs)、应用框架与软件分发层、应用层(科学计算、AI 数据科学)。到 2024 年,生态系统已经进化到包含专门用于 AI 推理的 TensorRT、用于数据科学的 RAPIDS 平台、支持容器化部署的 NGC 软件中心等先进组件。其兼容的深度学习框架也已经从 TensorFlow、PyTorch 拓展到了云平台环境,并支持包括 x86、ARM 在内的多种计算架构。这一生态的无可替代性使得 NVIDIA 发展成为全球人工智能基础设施建设的基础底座。

(三)应用层:打造跨行业、跨领域的深度融合机制

美国积极构建跨行业、跨领域的创新生态系统,打通科技成果“产学研用”链条,并通过定制化服务加速人工智能应用的市场适配。

(1)依托产学研合作园区打通科技成果产出链条。产学研合作园区是美国人工智能的创新枢纽。这些园区通常由高校、科研机构、政府及企业联合设立,是集孵化器、基础研究、风险投资、商业化于一体的创新生态系统。以斯坦福工业园为例,园区通过荣誉合作项目(HCP)、教师咨询服务和产业联盟计划等方式,将斯坦福大学的学术研究与园区企业的产业需求紧密结合起来。园区还通过租约条款,将入驻的公司严格限定为整体核心业务为从事研发的机构,以确保园区的创新氛

围。Alphabet、NVIDIA、思科、雅虎等众多塑造科技行业的公司均是在斯坦福园区成长起来的。

(2)培育开源生态。美国通过构建开放共享的创新生态,促进人工智能基础设施的跨行业、跨领域应用[14]。一方面,美国的科技企业和研究机构积极推动 AI 开源。美国企业与开源平台、开源社区已经形成较为完善的合作路径,谷歌开源的 TensorFlow 和 Meta 主导的 PyTorch 已成为全球深度学习领域的核心工具,被广泛应用于自动驾驶、智慧城市、计算机视觉等领域。另一方面,美国政府为 AI 开源提供保障。美国政府早在 2016 年出台了专门的源代码政策。其要求每个机构每年必须将其至少 20% 新定制的开发代码作为开源软件(OSS)向公众发布。2023 年,美国国会通过《保护开源软件法案》(Securing Open Source Software Act of 2023),要求发布一套包含政府、产业界与开源社区现有实践的统一风险评估框架,加强对开源软件安全的保障。美国开源软件安全倡议(Open Source Security Initiative,简称"OS3I")围绕开源安全展开了一系列努力,包括推动在公共部门设立开源项目办公室(OSPO)、积极协调公私合作,并启动了多个专注于开源软件安全审计和基础设施加固的项目。同时,作为更广泛的网络安全战略的一部分,美国政府(包括支持 OS3I 的机构)正积极倡导从内存不安全的语言(如 C/C + +)向内存安全语言(如 Rust)转型,以使从根本上减少软件漏洞,但这并非是一个"转换所有 C 代码"的单一项目。

(3)定制化服务构建商业价值闭环。定制化服务不仅是商业模式创新的关键,更是人工智能产业获取长期竞争力的重要途径。定制化服务的核心在于深入理解行业需求[15],并围绕特定应用场景提供个性化的人工智能解决方案。例如,亚马逊 AWS 的人工智能云服务专门为医疗、金融、制造等行业提供定制化计算资源及算法优化支持,使企业能够根据自身需求构建 AI 模型,从而提高应用效率。在零售领域,微软正通过其 Microsoft Cloud for Retail 平台,为零售商提供集成化的数据与人工智能解决方案。这些工具赋能企业利用自身数据,构建高度个性化的客户交互体验,从而提升服务水平和运营效率。在公共服务数字化方面,美国总务署(GSS)正在内部试点一个名为"GSA Ask"生成式 AI 聊天机器人项目。该工具旨在协助联邦雇员高效处理与管理政府办公、采购合同和信息技术设施相关的复杂查询与任务,是推动联邦政府内部运营现代化的重要尝试。

(四)政策层:强化人工智能基础设施的顶层设计与引导

(1)将人工智能基础设施建设纳入国家优先事项。美国不断强化人工智能基础设施领域的战略部署。2025 年 1 月,美国启动专门的人工智能基础设施建设计划"星际之门"。2025 年 7 月,美国发布的《赢得竞争:美国人工智能行动计划》对人工智能基础设施建设进行了系统性安排。分别是加速人工智能创新、建设人工智能基础设施和引领国际人工智能外交与安全。该计划指出,强大的人工智能基础设施是赢得全球人工之智能竞赛的根本前提,需要破除美国数据中心、芯片、能源系统面临的约束。在数据中心方面,美国为数据中心相关的建设活动设立新的"分类排除"条款。在芯片方面,致力于恢复美国的半导体制造业。在能源系统方面,专注于提高电网的韧性和可靠性,优先将稳定、零碳的先进基荷能源接入电网。这一系统化的人工智能基础设施发展战略为此后美国的长期技术领先和产业生态构建奠定了坚实基础。

(2)注重规范引导,保障安全与合规发展。2024 年 7 月欧盟《人工智能法案》是全球首个全面针对人工智能监管的综合性法规[16],其为人工智能的有效治理提供了宝贵经验。一是以风险为导向的精细化分级治理。法案将人工智能系统分为"不可接受风险""高风险""有限风险"和"极小风险"4 个等级。其中,法案明确禁止操控性技术、利用弱点、社会评分、情绪识别等处于"不可接受风险"类别的人工智能实践;法案要求用于关键基础设施管理、就业与人力资源管理、医疗设备等用途的"高风险"人工智能系统承担大部分合规义务;法案对"有限风险"类别的人工智能系统设定透明度义务,对"极小风险"类别的人工智能系统不施加额外的强制性法律义务。二是通过"监管沙盒"

等机制平衡监管与发展。为避免严格的法规扼杀创新,法案引入了“监管沙盒”机制。这一机制为企业提供了一个受控的法律“实验区”,允许企业在真实或模拟环境中测试、验证创新的 AI 系统,而无须担心因不合规而受到处罚。

三、我国人工智能基础设施建设现状与面临的主要问题

自 2017 年《新一代人工智能发展规划》将人工智能发展纳入国家战略规划以来,我国人工智能迅速发展,相继建成一批具有代表性的人工智能基础设施,其主要集中在算力网络、大模型及其研发生态两个方面。第一,依托鹏城实验室(PCL)和领先的人工智能企业推进全国一体化算力网络构建。2022 年 6 月,“中国算力网”(C2NET)正式上线一期工程“智算网络”,建成全国智能算力互联体系与人工智能开源开放平台。目前,我国已有 30 多个城市正在建设或规划建设智能计算中心或专门用于运行人工智能负载的基础设施[17]。截至 2024 年,我国提供算力服务的在用机架数达 830 万标准机架,算力总规模达到 246 EFLOPS[18]。截至 2025 年 6 月,我国智算规模已达到 788 EFLOPS,算力供给能力持续提升。第二,自主性大模型不断涌现,研发生态不断完善。中国政府大力支持国内人工智能大模型开发,自 2020 年以后,我国大模型迅速发展,拥有超三分之一的全球人工智能大型语言模型。阿里巴巴(Qwen)、深度求索(DeepSeek)和智谱 AI(GLM)已成为具有国际竞争力的大模型研发团队[19]。此外,我国不断强化对 Gitee 等国内开源软件社区、百度飞桨(PaddlePaddle)、华为昇思(MindSpore)等框架的支持,以此规避过于依赖 TensorFlow、PyTorch、LLaMA 等深度学习框架对我国人工智能研发生态造成的潜在风险。然而,我国人工智能基础设施的主要建设路径为政府战略性驱动加链接全球人工智能创新生态。在当前的发展情境下,我国人工智能基础设施的部署、应用仍面临一系列挑战,主要包括以下 4 个方面。

(一)基础层:资源整合能力薄弱,核心技术面临短板

一方面,我国算力、算法、数据资源要素的整合能力不足。当前,我国在整合算力、算法与数据这三大人工智能核心要素方面仍面临着显著挑战。尽管在算力基础设施和数据体量上具备规模优势,但资源普遍处于“碎片化”状态。特别是在数据领域,由于跨机构、跨行业的数据共享流通机制不畅,以及统一标准和合规路径的缺失,大量高价值行业数据被隔离在分散的“孤岛”中,甚至流失境外,导致真正可用于训练和优化 AI 模型的高质量中文数据集严重短缺。在算力层面,我国尚未完成通用算力、智能算力和超级算力的合理布局,不同算力主体的技术体系、基础架构、调用接口不统一,多层次、跨区域的算力网络尚未完全形成,西部算力基础设施整体空置率超过 50%[20]。此外,我国国产芯片与算法之间的适配性较低,导致大量的算力浪费。

另一方面,我国在图形处理单元(GPU)的张量处理器(TPU)自主性和创新上面临技术短板。大模型分布式训练是典型的计算密集型和通信密集型场景。长期以来,依托领先的芯片架构和深度耦合的 CUDA 生态优势,NVIDIA、AMD、Intel 等国际巨头几乎垄断了全球 GPU 市场。近年来,我国虽涌现出以寒武纪、景嘉微、华为昇腾等为代表的自主 AI 与图形计算芯片企业,但其产品在综合竞争力上仍难以与国际巨头直接抗衡。据行业分析,在硬件性能,尤其在软件生态(如开发生态 CUDA 的替代性)和产业链成熟度方面,国内 GPU 及 AI 加速器产品与以 NVIDIA 为代表的全球领先水平存在 2 ~3 年的代际差距。与此同时,在专注人工智能训练的张量处理器(TPU)领域,以谷歌 TPU 为标杆,我国同类 AI 加速器在峰值算力、能效比与软件栈的易用性和完善度以及参与制定国际行业规范的话语权等方面,依然面临显著挑战。核心差距不仅在于单芯片性能,更在于构建围绕自身硬件的、繁荣的全球开发者与应用生态。

(二)模型层:高质量大模型缺乏,研发生态有待完善

我国大模型的“质”与“量”与美国仍存在较大差距。根据斯坦福大学发布的《2024 年人工智能指数报告》的数据显示,2023 年美国发布 61 个著名 AI

模型和109 个基础模型,我国的数量产出则分别为5个和20 个[21]。尽管我国已涌现出一批自研大模型,但部分模型在研发过程中仍不同程度地参考或依赖于国际主流模型架构(如 Transformer)及开源生态(如 GitHub 平台)。在模型的原创性、体系化创新和综合性能上,与国际顶尖水平仍存在差距。此外,相较于美国少数大型科技企业与超大规模云提供商联合聚拢资源的开发模式而言,我国的人工智能大模型开发环境较为分散,目前约有 50 家企业参与大模型开发,这种分散性在制约资源聚合的同时也可能造成无序开发。

我国的大模型研发生态仍有待进一步完善。我国的开源托管平台在大模型数量、用户数量上与 GitHub、Hugging Face 存在较大差距。数据显示,GitHub 拥有 1.5 亿名以上开发人员、400 个开源组织、4.2 亿个多存储库。截至 2025 年 7 月,Hugging Face 共享 1 559 464 个模型、346 999 个数据集。相较之下,我国知名的开源托管平台如 Gitee 的用户数量仅为 1 350 万名、开源组织数量为 40 万个,魔搭社区(ModelScope)的开发者、优质模型数量则分别约为 560 万名和 5 500 个。需要注意的是,我国的开源托管平台之间并不能实现共享,存在“数量多,聚合低,同质化高”的缺陷。全链条的模型开发工具与高质量开源平台的不足,对我国人工智能发展的限制作用可能在未来被进一步放大,成为“卡脖子”的一环。

(三)应用层:基础研究和产业技术的信息化基础设施不够完善,头部企业作用有限

一是我国的开放科学体系尚不发达。开放科学涵盖了科学知识与开放数据、科学基础设施与平台开放、公众参与开放等多个层面[22]。当前,我国在科研投入和总量方面取得了显著成绩,但在开放科学生态体系的建设上仍然相对匮乏。首先,科研数据共享的透明度和规范性存在不足。我国已逐步建立了科研数据管理和共享平台,但整体上数据的开放水平参差不齐,数据和存储仍未形成统一的规范,阻碍了数据的交换和使用。其次,国内开放出版和开放获取进展滞后。国内权威数据平台与期刊编辑部之间在开放问题上尚未形成有效协同,整体开放获取的比例仍然有限,且往往面临着潜在的知识产权风险和法律纠纷[23]。最后,开放科学基础设施建设不足,缺乏国家级的人工智能开放科学基础设施,限制了不同平台、机构、地区间人工智能资源的流动与利用。

二是我国科技企业在推动人工智能基础设施建设的作用有限。目前,我国人工智能企业更多地是将资源集中于应用层面,如智能终端、互联网平台和垂直行业解决方案,在底层算法框架、通用大模型研发和高端算力基础设施建设上相对薄弱。数据显示,我国应用层企业占我国所有人工智能骨干企业的61.47%,基础层和技术层企业的占比则分别为 9.93%、28.6%[24]。此外,我国人工智能企业的行业标准引领作用不足,表现在我国人工智能企业在国际电信联盟(ITU)、国际标准化组织(ISO)等国际标准制定组织中的参与度和转化率不高,从而限制我国企业在全球形成广泛的技术扩散和产业联盟。

(四)政策层:政策协调有待加强,安全监管有待强化

国家和地方层面均出台了大量的人工智能政策。这些政策的内容可以总结为6 项类别,分别是试点示范、学校教育、技术创新、经济发展、智能应用和人力资源[25]。通过分析现有政策,可以发现支撑人工智能基础设施发展的政策体系仍存在以下不足。

第一,区域差距有待弥合。就政策数量而言,政策发布在地理上存在明显的不均衡性,政策数量排名前五的省份均为沿海地区。这种数量上的差距揭示了中西部地区在人工智能基础设施的建设意愿、建设能力、建设资源上的不足。中西部与经济发达省份的数字鸿沟可能会进一步加剧,欠发达地区由于资金匮乏、技术欠缺等在其中难以充分受益,接入型、使用型、能力型数字鸿沟将进一步扩大[26]。

第二,政策工具结构失衡。当前的政策体系在工具选择上呈现出明显的偏好,即过度侧重于供给侧工具(如直接资助研发、建设算力基础设施)与环境侧工具(如制定技术规划、设立法规框

架),需求侧工具(如公共采购、应用示范补贴、消费者激励)的应用则相对匮乏与薄弱。人工智能基础设施的建设与技术的研发,在一定程度上脱离了真实市场场景的有效牵引,可能面临与产业需求脱节的风险,难以形成“市场需求驱动技术创新—技术升级创造更大市场”的良性循环,最终可能影响整个产业生态的可持续发展[27]。

第三,安全治理体系滞后。尽管 2023 年 8 月实施的《生成人工智能服务管理暂行办法》填补了生成式 AI 领域的监管空白,但其治理框架在应对快速演进的技术风险时已显现出多重不足。其一,监管覆盖存在盲区,当前法规主要规制面向公众的终端服务,对未公开的底层模型及其研发过程的监管覆盖不足。其二,缺乏差异化的行业指引,未能针对金融、医疗、司法等高风险领域的特殊需求制定分级分类的合规标准。其三,也是最为关键的,是对人工智能基础设施的风险防控关注缺位。随着算力中心、大规模数据平台等新型基础设施的广泛部署,其本身已成为网络攻击的高价值目标,面临着供应链安全、模型泄露、数据投毒等独特威胁,而现行法规在此层面的风险预警与管控措施尚属空白,构成了潜在的系统性安全短板。

四、促进我国人工智能基础设施建设的启示与建议

我国应瞄准人工智能基础设施建设的痛点,充分吸收人工智能基础设施建设的先进国际经验,推动人工智能基础设施多维度、全方位部署,具体包括以下 4 个方面。

(一)基础层:强化资源整合流通,夯实基础支撑

(1)推动核心技术突破。一是,聚焦人工智能基础性、前沿性、通用性技术,设定人工智能技术清单与风险清单,根据紧迫性、重要性进行分级分类,并定期审查与更新。二是,培养人工智能基础设施建设的人才。鼓励百度、阿里巴巴、腾讯、华为等头部科技企业与科研院所、高校合作共同开发人工智能领域的实验课程、技术培训和人才基地。三是,建立人工智能基础设施技术突破转向资金。通过财政拨款、税收优惠、专项资金、科学基金、政府采购、社会资本等,建立竞争性支持和稳定性支持相结合的基础研究投入机制,鼓励金融资源流向人工智能技术创新领域。

(2)加快推进算力、算法、数据资源互联互通。一是,加快资源整合和调度。加快整合互联网企业、电信运营商等算力持有者的资源,推进以应用为导向的全国一体化算力网络体系建设,提高算力资源的可用性和共享性。二是,构建统一的算力互联互通标准。在硬件层定义统一的硬件接口、在网络层制定专门面向人工智能计算的通信协议、在软件层推广标准化的容器技术和通用的应用调度 API。三是,强化数据统筹。充分发挥国家数据局在数据方面的统筹作用,建设全国性的数据交换开放平台与数据训练集,盘活国家科学数据中心和国家资源库等科技资源,鼓励企业和科研机构共享高质量数据集,打造跨部门、跨领域、跨系统间互联互通的数据生态。

(3)重视人工智能基础设施的“绿色属性”[28]。一是,将与人工智能基础设施相关的能源需求预测纳入国家电网脱碳规划和投资。建立最佳实践指标来评估和报告整个人工智能价值链的碳排放、能源消耗以及电网的容量和碳强度。二是,加强对数据中心的碳排放管理。可以效仿欧盟的做法,考虑对不同规模的数据中心提供商设置一定的能源效率和碳排放门槛。如 2024 年 7 月我国发布的《数据中心绿色低碳发展专项行动计划》就明确要求到 2025 年底全国数据中心平均电能利用效率降至 1.5 以下。三是,促进私营部门在清洁能源方面的技术创新。谷歌和加州大学伯克利分校的研究表明,采用适当的算法、定制硬件和节能的云数据中心能够使大型语言模型(LLM)的碳足迹减少到原来的 1/1 000 ~ 1/100[29]。建议建立促进清洁能源进步的商业化政策,通过贷款计划、差价合约(CFD)、电力购买协议(PPA)、承购协议等方式推动私营企业开展绿氢、模块化核反应堆、车网互动、地热等方面的创新与研发。

(二)模型层:提升模型质量,完善研发生态

(1)提升大模型国际影响力。大模型的国际影响力来源于模型质量,而质量体现于模型在泛

化性能、安全可控性和跨领域适用性等方面的综合表现。一是,要增强大模型的广泛适配能力。这一能力的增强可以通过3 种方式实现,包括让模型在特定领域的海量文本进行深度学习的领域自适应预训练[30]、采用“混合推理”框架进行指令微调[31]利用检索增强生成技术使模型能动态访问外部最新信息[32]。二是,要增强大模型的多模态与代理能力。随着人工智能从单一任务向综合认知发展,多模态处理能力成为模型竞争力的重要标志。通过跨模态内容的高效整合、理解与生成,围绕医疗影像诊断、工业质检、教育互动等关键行业和领域,构建多模态任务数据集并进行定向优化。

(2)构建全栈式的人工智能开源平台。一是,强化我国现有开源平台之间的协调。发挥好新一代人工智能产业技术创新战略联盟(AITISA)、中国人工智能产业发展联盟(AIIA)、开源中国(OSCHINA)等行业组织在推动我国人工智能基础设施建设与协同方面的能力,积极探索 GitHub、Gitee、GitCode、ModelScope、开放传神(Open CSG)等开源平台建设跨平台同步功能[33]。二是,提升我国开源平台在国际层面人工智能基础设施标准、规范方面的参与能力。积极参与国际人工智能标准化活动,推动中国在人工智能基础设施领域的技术标准、测试规范、数据隐私保护等方面形成“先行者优势”。三是,拓展高质量资源。通过共同设立联合实验室、技术攻关项目、共建人工智能基础设施项目、资金合作等方式加强与欧盟、东盟、全球顶尖实验室之间的人工智能数据集、知识、技术的共享与联动。

(三)应用层:拓展应用场景,强化企业作用

(1)拓展人工智能基础设施应用场景,推动技术与市场深度融合。一是,制定行业导向性政策,打通跨行业应用场景。重点应着眼于消除行业间的壁垒、促进数据流通与共享、打破技术应用的瓶颈。二是,强化人工智能技术在传统行业中的融合、应用。围绕数字产业化与产业数字化的导向[34],重点推广成熟的人工智能软件、硬件产品,推进传统制造业全流程自动化、数字化、智能化。三是,建设人工智能行业应用示范区,推动技术验证与应用规模化。重点选择智能制造、智慧城市、健康医疗等代表性的领域作为技术验证和应用的先行区,推动标准化技术在区域内的集成应用和推广示范。四是,打造开放创新氛围。从开放科学角度引导国内重大科研基础设施、研究成果共享[35]。同时,完善开放科学的体制机制,考虑将开源纳入学术评价体系,打造开放激励机制,鼓励科研人员积极参与开源项目、开源社区建设开发。

(2)进一步发挥国内人工智能头部企业在开发、应用人工智能基础设施方面的作用。一是,推动人工智能头部企业成为计算平台、云服务平台、算法框架等人工智能基础设施的优质提供者。建议引导头部企业加强对数据处理、算法研发等全链条技术体系的投资与融合建设,以企业为支点撬动整个产业链的良性发展与创新。二是,提升头部科技企业的投资能力和协同合作水平。拓展大型科技企业的金融角色,鼓励大型科技企业设立人工智能专项投资基金。同时探索大型科技企业提供低息贷款的发展路径,帮助中小企业建设和部署人工智能基础设施。三是,发挥示范引领作用,推动行业标准化和市场化进程。例如,华为可以依托在 5G、智能硬件和边缘计算领域的技术积累提供智能城市建设的技术方案,从而推动建立标准化的人工智能技术应用“标杆”,为其他企业和行业提供操作指南和实施框架。

(四)政策层:增强战略统筹,强化安全监管

(1)构建国家级人工智能基础设施发展战略。以新质生产力为抓手,将新型人工智能基础设施与数字中国、开放科学、可持续发展等国家战略目标相结合,形成统一的人工智能基础设施建设的短期、中期、长期战略规划。推动成立负责全国性人工智能基础设施的专门机构或小组,在顶层设计层面进行统一部署安排。同时,厘清各级政府部门的主要职能与定位,强化中央和地方关于人工智能基础设施建设方面的协调与联动,充分发挥大型科技企业、社会组织在行业内的能动性作用,打造多元协同的发展与治理格局[36]。

(2)强化多元规则构建,就人工智能基础设施的开发、部署、运用的全链条进行监管。首先,完

善人工智能基础设施相关技术的安全标准与法律法规。借鉴国际经验,探索动态的风险分级机制,要求企业在人工智能技术研发和应用过程中进行全面的安全性评估与风险预判,并对风险内容进行定期报告。其次,强化对人工智能基础设施开发和运营的市场监管,制定相关政策避免市场垄断和不正当竞争出现。最后,加强对数据市场的监管,尤其是在涉及个人隐私、国家安全等敏感信息时,必须确保数据采集、存储、处理和使用的全过程符合相关法律法规,确保数据的合规流通与安全共享。同时,坚持定期开展安全审计、漏洞评估与代码审查,强化对于开源基础模型的全面风险评估、成本收益核算。

五、结语

人工智能基础设施构成了数字经济时代发展的核心支撑,其包括基础层、模型层、应用层与政策层四个维度。坚实的基础技术、良性的研发生态、完整的应用机制、完善的顶层设计共同构成了人工智能基础设施部署、应用的肥沃“土壤”。自2017年以来,我国人工智能进入系统性快速发展阶段,全国一体化算力网络加速部署,国内自主性大模型不断涌现,建成一大批具有代表性的人工智能基础设施。值得注意的是,我国人工智能基础设施的建设仍然面临基础资源整合能力不足、高质量大模型缺乏、应用机制不完善、政策统筹不健全等问题。我国应立足发展实践,充分借鉴国际先进经验,重点聚焦基础层资源整合、模型层生态完善、应用层场景拓展、政策层战略统筹,推进建设具有国际竞争力的人工智能基础设施,推动人工智能基础设施等数字公共产品开放协作可持续建设发展[37],在不确定的地缘政治环境中为我国人工智能的稳健、快速发展构建牢固的战略支撑。

参考文献:

[1]中华人民共和国工业和信息化部. 工业和信息化部等十一部门关于推动新型信息基础设施协调发展有关事项的通知[EB/OL]. (2024-04-08)[2024-10-10]. https://www.miit.gov.cn/jgsj/txs/wjfb/art/2024/art_fd2f0e91ff0d45209ae74744d58b6b50.html.

[2]RedHat. AI infrastructure explained[EB/OL]. (2024-04-08)[2025-05-31]. https://www.redhat.com/en/topics/ai/ai-infrastructure-explained.

[3]RCR. AI infrastructure—mapping the next economic revolution[EB/OL]. (2025-04-11)[2025-05-25]. https://www.rcrwireless.com/20250411/ai-infrastructure/ai-infra-revolution.

[4]HUA H, LI Y, WANG T, et al. Edge computing with artificial intelligence: a machine learning perspective[J]. ACM computing surveys, 2023, 55(9): 1-35.

[5]FATTORINI L, MASLEJ N, PERRAULT R, et al. The global AI vibrancy tool[EB/OL]. (2024-11-14)[2025-03-25]. https://hai.stanford.edu/research/the-global-ai-vibrancy-tool-2024.

[6]ZENG Y, LU E, GUAN X, et al. AI governance international evaluation index (AGILE Index)[R]. Beijing: Center for Long-term Artificial Intelligence (CLAI), 2025: 13-15.

[7]ARTIFICAL ANALYSIS. State of AI: China-artificial analysis Q1 2025[EB/OL]. (2025-02-06)[2025-03-25]. https://artificialanalysis.ai/downloads/china-report/2025/Artificial-Analysis-State-of-AI-China-Q1-2025.pdf.

[8]NSF. Strengthening and democratizing the U.S. artificial Intelligence innovation ecosystem: an implementation plan for a national artificial intelligence research resource[EB/OL][2025-03-25]. https://nsf-gov-resources.nsf.gov/2023-10/NAIRR-TF-Final-Report-2023.pdf.

[9]NTIS. Supercomputers: government plans and policies[R]. Washington, DC: U.S. Government Printing Office, 1986: 11-17.

[10]BRITNEY N. The top 10 countries in the world by computing power Britney Nguyen[EB/OL]. (2025-01-06)[2025-03-25]. https://qz.com/top-10-countries-world-compute-power-supercomputing-ai-1851727906.

[11]CRAWFORD K. The atlas of AI: power, politics, and the planetary costs of artificial intelligence[M]. New Haven, CT: Yale University Press, 2021: 41.

[12]MIRUNA M. New data center sustainability reporting obligations introduced through the European commission delegated regulation (EU) 2024/1364[EB/OL]. (2024-09-09)[2024-10-28]. https://www.dlapiper.com/en/insights/publications/2024/09/new-data-centre-sustainability-reporting-obligations-introduced-through-eu-20241364-regulation.

[13]MASLEJ N, FATTORINI L, PERRAULT R, et al. Artificial intelligence index report 2025[R]. Stanford, CA: Institute for Human-Centered AI, Stanford University, 2025: 4-49.

[14]隆云滔,王晓明,顾荣,等. 国际开源发展经验及其对

我国开源创新体系建设的启示[J]. 中国科学院院刊, 2021,36(12):1497-1505.
[15]郁建兴,刘宇轩, 吴超. 人工智能大模型的变革与治理[J]. 中国行政管理, 2023(4): 6-13.
[16]FLI. High-level summary of the AI Act[EB/OL]. (2024-02-17)[2024-08-15]. https://artificialintelligenceact. eu/high-level-summary/.
[17]REBECCA A. China's AI development model in an era of technological deglobalization [EB/OL]. (2024-05-02)[2024-10-28]. https://merics. org/en/report/chinas-ai-development-model-era-technological-deglobalization.
[18]中国信息通信研究院. 先进计算暨算力发展指数蓝皮书(2024 年)[R]. 北京: 中国信息通信研究院, 2025: 4-5.
[19]NATHAN B, ALEX C. The state of Chinese AI:SOTA research, chips, sanctions, and loopholes[EB/OL]. (2024-07-18) [2024-11-05]. https://press. airstreet. com/p/the-state-of-chinese-ai.
[20]洪涛,程乐. 全国算力体系一体化建设的五大问题及治理对策[J]. 中国科学院院刊, 2024,39(12):2086-2095.
[21]MASLEJ N, FATTORINI L, PERRAULT R, et al. Artificial intelligence index report 2024[R]. Stanford, CA: Institute for Human-Centered AI, Stanford University, 2024: 4-13.
[22]FECHER B, FRIESIKE S. Open science: one term, five schools of thought[M]. Cham: Springer International Publishing, 2013:17-47.
[23]王哲,隆云滔. 中美知识产权博弈与中国知识产权制度建设[J]. 东北财经大学学报,2023(4):19-29.
[24]前瞻产业研究院. 2025 年中国人工智能行业全景图谱[R]. 北京:前瞻产业研究院,2025:4-10.
[25]MNEKGIR H J. The US-Chinese race in artificial intelligence challenges and opportunities[J]. Russian law journal, 2023, 11(3):2578-96.
[26]王驰,曹劲松. 数字新型基础设施建设下的安全风险及其治理[J]. 江苏社会科学, 2021(5):88-99,242-243.
[27]宋琪, 谷灏. 政策"工具—功能"视角下人工智能产业央地政策研究[J]. 科学与社会, 2022, 12(1): 84-102.
[28]ROBBINS S, VAN WYNSBERGHE A. Our new artificial intelligence infrastructure: becoming locked into an unsustainable future[J]. Sustainability, 2022, 14(8): 4829.
[29]PATTERSON D, GONZALEZ J, LE Q, et al. Carbon emissions and large neural network training[J/OL]. arXiv preprint arXiv, 2021, 2104. 10350:1-2. (2021-04-23)[2024-10-28]. https://arxiv. org/abs/2104. 10350.
[30]KE Z, MING Y, JOTY S. Adaptation of large language models[R]// Proceedings of the 2025 Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies. Albuquerque: Association for Computational Linguistics, 2025: 1-8.
[31]XU F, HAO Q, SHAO C, et al. Toward large reasoning models: a survey of reinforced reasoning with large language models[J]. Patterns,2025, 6(10):16.
[32]ZHENG M, CHEN H, GUO T, et al. Enhancing large language models through adaptive tokenizers [J]. Advances in neural information processing systems, 2024, 37: 113545-113568.
[33]王哲,薛澜. 大模型开源创新公地:历史演进、价值逻辑与中国叙事[J]. 探索与争鸣,2025(7):128-137,179.
[34]中国信息通信研究院. 全球数字经济白皮书[R]. 北京: 中国信息通信研究院, 2023:1-4.
[35]李秋实,刘欢,张顺顺. 生成式人工智能变革下开放科学的发展机遇与推进策略[J]. 图书情报工作, 2024, 68(18): 13-22.
[36]隆云滔, 王哲, 许哲平,等. 推动国家资助科研项目成果开源开放的国际经验借鉴及思考[J]. 中国科学院院刊, 2024, 39(1): 152-162.
[37]隆云滔,林靖玲,刘海波,等. 数字公共产品开放协作治理机制研究:基于国际经验分析的视角[J]. 中国软科学, 2025 (1):65-76.

(本文责编:润　泽)